Manipulation arguments and the freedom to do otherwise

Manipulation Arguments and the Freedom to do Otherwise

Patrick Todd

The University of Edinburgh

Abstract

I provide a manipulation-style argument against classical compatibilism – the claim that freedom to do otherwise is consistent with determinism. My question is simple: if Diana (the designer) really gave Ernie (the designed) free will, why isn't she worried that he won't use it precisely as she would like? Diana's non-nervousness, I argue, indicates Ernie's non-freedom. Arguably, the intuition that Ernie lacks freedom to otherwise is stronger than the direct intuition that he is simply not responsible; this result highlights the importance of the denial of the principle of alternative possibilities for compatibilist theories of responsibility. Along the way, I clarify the dialectical role and structure of "manipulation arguments", and compare the manipulation argument I develop with the more familiar Consequence Argument. I contend that the two arguments are importantly mutually supporting and reinforcing. The result: classical compatibilists should be nervous – and if PAP is true, all compatibilists should be nervous.

Introduction

In contemporary debates about moral responsibility and determinism, one often encounters what have come to be called *manipulation* scenarios – roughly, scenarios in which a given agent's behaviour has been secured in advance by powerful manipulators or designers working "behind the scenes". Intuitively, so the thought goes, compatibilists will have to say that such agents may indeed be morally responsible for what they do. But (so the proponents of the arguments maintain) it is, intuitively, false that these agents may be responsible; so compatibilism is false. The relevant scenarios are meant to elicit an intuition of non-responsibility. Broadly speaking, however, this has been an intuition compatibilists have either not shared or have otherwise wished to explain away. Such has been, thus far, the debate concerning so-called "manipulation arguments".

Manipulation arguments, then, have primarily targeted moral responsibility – and they have targeted it *directly*. In this paper, however, I wish to point out that proponents of such arguments may also target moral responsibility *indirectly*, by first targeting the *freedom to do otherwise*, and then by appeal to the traditional requirement, enshrined in the "principle of alternative possibilities" (PAP), that freedom to do otherwise is necessary for moral responsibility. I believe that this is an underappreciated strategy in the compatibility debates. As I hope to show, the intuition that the relevant (manipulated) agent lacks the freedom to do otherwise is arguably *stronger* than the (direct) intuition that the agent is simply not responsible. In this paper, then, I will offer a manipulation argument against *classical compatibilism* – the claim the freedom to do otherwise is compatible with causal determinism. Together with the principle of alternative possibilities, it would follow that moral responsibility is incompatible with determinism – but I will not argue for any such principle here.[1] I simply wish to highlight the *availability* of such an argument.[2] Moreover, the failure of classical compatibilism, if it is a failure, is, of course, of independent interest, apart from the (putative) connection to moral responsibility.

Thankfully, the argument to come is simple – and brief. It comes in the form of a story, and it begins with a familiar goddess named Diana. I call it the "Why aren't you worried?" argument, for reasons that will become clear momentarily. I then discuss the relationship between this argument and two more traditional arguments in the free will/moral responsibility debates: (1) the "Consequence Argument" (for the incompatibility of the freedom to do otherwise and determinism) and (2) the "Direct Argument" (for the incompatibility of moral responsibility and determinism). Along the way, I attempt to clarify the dialectical role and structure of "manipulation" arguments in the relevant debates.

The Zygote Argument

---

[1] PAP, of course, has famously been called into question by the so-called "Frankfurt examples". But the jury on such examples is still out; see, e.g., Franklin 2011, Capes 2014, and Swenson 2015 for recent replies to the Frankfurt cases. The present paper highlights the importance of the denial of PAP for compatibilist theories of moral responsibility.

[2] As I discovered after this paper was accepted, the availability of this strategy has indeed already been highlighted by Cohen 2015. Cohen's paper, however, focuses on Pereboom's "Four Case" argument and the classical compatibilism of Kadri Vihvelin, and, in any case, does not appeal to the considerations about "nervousness" central to my argument below. We thus employ substantially different resources to arrive at similar conclusions. Naturally, however, I am sympathetic with much of what Cohen says.

Consider the following scenario described by Alfred Mele, which we can call the "Diana scenario":

> Diana [a goddess with special powers] creates a zygote Z in Mary. She combines Z's atoms as she does because she wants a certain event E to occur thirty years later. From her knowledge of the state of the universe just prior to her creating Z and the laws of nature of her deterministic universe, she deduces that a zygote with precisely Z's constitution located in Mary will develop into an ideally self-controlled agent [Ernie] who, in thirty years, will judge, on the basis of rational deliberation, that it is best to A and will A on the basis of that judgment, thereby bringing about E.[3]

In short, we're asked to imagine that Diana (1) knows the (deterministic) laws of nature, (2) has the unilateral power to secure various sets of "initial conditions", and (3) the computational power to deduce what happens in the future, given those conditions. We then suppose that she wants there to be a given agent in 30 years that performs various actions, and thus secures the relevant set of initial conditions so as to guarantee such an agent performing those very actions. We're then asked to consider whether any such agent could fairly be blamed for performing them, in case the actions Diana wished the agent to perform would ordinarily be morally objectionable. It is expected that we would conclude that no such agent could fairly be blamed for performing those actions. Compatibilists, however, will have to suppose otherwise – for there is (plausibly) no difference between someone's actions having been "set up" in the relevant way by *Diana* or instead by (otherwise similar) blind natural causes. In short, if you think that Ernie could not be to blame in the Diana scenario, then you must think that no one could be to blame in *any* deterministic scenario. Accordingly, you must think that compatibilism about moral responsibility and determinism is false.[4]

Why isn't Diana worried?

[3] Mele 2006: 188.

[4] For a structurally similar (and highly influential) manipulation argument for incompatibilism, see Pereboom's "Four Case" argument (the most recent statement of which is in Pereboom 2014: Ch. 4).

Thus proceeds the *direct* manipulation argument for incompatibilism concerning responsibility and determinism. But consider now the following modifications to the "Diana scenario". We are, somehow, introduced to a powerful goddess, Diana, about whom, initially, we're told nothing besides that she has extraordinary powers of creation. We discover that Diana wishes for there to be an agent who, in 30 years, faces a decision between lying and telling the truth. And she wishes this agent (Ernie) to be *torn* about this decision; that is, she wishes Ernie to have and to see reasons on both sides of lying and telling the truth, in such a way that lying and telling the truth are both fully consistent with Ernie's character. Further, and crucially, Diana wishes it to come to pass that Ernie lies – but had it *within his power* to tell the truth. That is, Diana wishes it to come to pass that though it is *up to* Ernie whether to lie or to tell the truth, Ernie in fact lies. For Diana, it is important that Ernie be *free* not to lie – but also that he lies. Otherwise, she says, her plan is a failure. Accordingly, she creates a set of initial conditions, presses play, and, as witnesses to her plans and activities, we wait. For 29 years and 364 days.

It comes to pass that there is a man named Ernie who feels torn about lying and telling the truth. And, we recall, it is highly important for Diana's plans that Ernie lie. The moment of truth is approaching. Ernie is deliberating. And we've been told – and have been assuming – that, as promised, Diana has given Ernie *free will* with respect to lying and telling the truth. Some of us begin thinking: if I were Diana, I'd be getting nervous. After all, Diana *wants* him to lie, but has left him *free* not to lie, and he's *torn* about lying.

But then we notice something curious. Diana seems perfectly relaxed. In fact, she is tweedling her thumbs (in her celestial throne room). We're struck by her total lack of concern. Thus begins the following exchange:

Us: Diana, isn't it important to you that Ernie lies here in a few minutes?

Diana: Yes.

Us: You realize that he's currently deliberating about doing just that? And that he's feeling awfully torn about what to do?

Diana: Yes.

Us: And didn't you say that you've left it up to Ernie whether to lie or tell the truth? That is, didn't you say that it would be up to him which to do?

Diana: Yes.

Us: Well, if it is really up to him, aren't you in the least bit *worried* that he's in fact going to tell the truth?

Diana: No.

Us: Why on earth not?

Diana: Because I am entirely certain that he will lie.

Us: But I thought you gave him free will – so couldn't he refrain?

Diana: I did give him free will, and yes, he has the power to refrain. But I'm entirely certain that he won't.

Us: What, do you have a crystal ball or something?

Diana: No, I don't have a crystal ball. And I have no magical powers of prevision. And nor am I somehow "outside time". I'm just a goddess, not *God*.

Us: Then how can you be so certain that Ernie's going to lie, if he has the power to do otherwise, and sees every reason in the world not to?

Diana: Because someone can be free not to do something, even if there is no objective chance whatsoever that he won't. And that's what I've arranged. Ernie is causally determined to lie, but is perfectly free not to. Neat trick, isn't it?

And we walk off thinking: this is, indeed, a trick – if there ever was one.

The point is simple. Genuine freedom, we think, should worry even someone like Diana. But this is not so, on classical compatibilism. Diana might have given Ernie free will, but might nevertheless have *perfectly planned* everything he ever does, and be appropriately *absolutely certain* that he will do only and everything she expects for him to do. But this result seems unacceptable. If Diana's plan involved Ernie's having free will with respect to lying or telling the truth, then it should have involved her *taking a risk*, viz., the risk that he would tell the truth. But she did not, in the above story, take any risk whatsoever. Accordingly, it seems plain that she did not – contrary to her claim – indeed give Ernie free will. Indeed, it seems *farcical* for Diana to claim that she gave Ernie free will concerning whether to lie or tell the truth, while sitting back and relaxing, having arranged matters so as to causally guarantee that he shall lie. Accordingly, whatever conditions the classical compatibilist identifies as being sufficient for the power to do otherwise are not sufficient, for Ernie (evidently) lacks any such power, but may perfectly well meet those conditions. More generally, any analysis of "could have done otherwise" that vindicates Diana's responses in the above exchange is thereby a defective analysis, for Diana's responses are defective. Classical compatibilism – the position she represents as being the case – is therefore defective.

To elaborate. It seems intuitive that, in the absence of a crystal ball, to believe that someone will perform an action he is free to refrain from performing, and torn about performing, is, necessarily, to risk being proved wrong. If his doing it is really up to him, then you'd have to be going out on a limb — at least somewhat — in believing that he'll do it. That seems to be implied by our concept of freedom (to do otherwise); freedom, if we had it, would always involve the ability to prove wrong anyone who believed that we would do one thing rather than another. But if free will is compatible with determinism, this isn't so, for it could be true that someone will freely do something, but one would not at all have to risk error in so believing, since one's belief could be based on causal factors and laws that entail that the action occurs. And this seems odd. This is, I believe, a result the (modified) "Diana scenario" helpfully brings out – and brings out in a forceful way. If our belief that someone will do something is based on present causal factors and laws that *guarantee* that person's action, then we take *no risk at all* in believing that the person will perform that action. There thus seems to be a deep tension in regarding that person as *free to refrain* from performing that action, for (again, in absence of a crystal ball) to believe that someone will perform an action he is free not to perform (and, we might add, has motive and reason not to perform) *is* to risk

being proved wrong. *Ex hypothesi*, however, we take no such risk. The claim of freedom, then, seems questionable at best.

I believe that these reflections allow us to see part of what is (arguably) so troubling about manipulation scenarios – or about taking ourselves to be like the subjects in such scenarios. We think that if we had genuine freedom, no one could just sit back and relax in perfect confidence that we will do only and exactly what she wants. But that's the picture we get with Diana: she can just sit back and relax. Real freedom would make her nervous. If Diana isn't nervous, then, it is because she didn't give Ernie real freedom – that is, she didn't really make it genuinely up to him what he does or becomes.

Two Objections

Here is the first:

> You say that Ernie's freedom ought to engender Diana's nervousness. But in the history of western theism, we encounter a long tradition of thought according to which God infallibly knows exactly what we're going to do, and yet we retain freedom to do otherwise. Many people take this point of view entirely seriously. It is not obvious, then, that, *necessarily*, freedom to do otherwise ought to engender just *anyone's* nervousness.

Granted. But this is why I have emphasized that, *lacking a crystal ball*, Ernie's freedom ought to imply Diana's nervousness. God, however, has a crystal ball – that is, with God, I admit, all bets are off.[5] Recall this bit of the exchange:

> Us: But I thought you gave him free will – so couldn't he refrain?
>
> Diana: I did give him free will, and yes, he has the power to refrain. But I'm entirely certain that he won't.
>
> Us: What, do you have a crystal ball or something?

[5] However, if it were to become clear that God *doesn't* have a crystal ball, but simply knows the future by determining it, then I would wish to say the same thing about God as I do about Diana; we could repeat the same argument in the case of God as I have given above. I thank the editor of this journal for raising this point.

Note precisely how *natural* this question is in this context. When Diana says that she is entirely certain that Ernie will lie, despite Ernie's being torn about lying, and free not to lie, precisely the natural question here is: what, can you simply see the future? Now imagine that Diana's response had instead been:

> Diana: Yes. Or I have what can only be explained to you, a mere mortal, as the possession of a crystal ball.

If we were prepared to believe Diana on this score, then I admit that a natural extension of our conversation would be:

> Us: Oh. Wow. That is amazing.

But we would not, I believe, immediately call into question Ernie's freedom, if we were assuming that the *grounds* of Diana's confidence (somehow) come from the future itself. (At most, the philosophers amongst us would start raising hackles about the possibility and presuppositions of Diana's alleged crystal ball.) Further, just as we may note precisely how natural the above question sounds, we may note precisely how *unnatural* the following question sounds:

> Diana: I did give him free will, and yes, he has the power to refrain. But I'm entirely certain that he won't.
>
> Us: What, did you put in place causes sufficient to determine that Ernie shall lie, but which nevertheless leave him free to refrain from lying?

This is, I submit, a question no one but a philosopher with a theory would ever think to ask.

Here is the second objection:

> In ordinary life, we often know what people are going to do, mostly because we know these people intimately – we know their characters and dispositions and the like. But we do not thereby assume that, strictly speaking, they *cannot* do

> otherwise than what we know they're going to do. For instance, I've known my significant other here for 25 years. And I know that, here in a few moments, she's going to order vanilla ice cream from the counter; that's her favorite, and that's what she always does. But of course she's *free* not to order vanilla, and instead to order chocolate.

I am prepared to admit that this is a natural thing to say. But this is why I have emphasized that Ernie is *torn* with respect to lying and telling the truth, and that both lying and telling the truth are fully consistent with Ernie's character. That is, what we are considering here is *not* a case in which Ernie's decision to lie is – as we would ordinarily say – predictable on grounds of Ernie's settled dispositions. Suppose, however, that this aspect of the setup of the story were removed. And now imagine:

> Us: But I thought you gave him free will – so couldn't he refrain?
>
> Diana: I did give him free will, and yes, he has the power to refrain. But I'm entirely certain that he won't.
>
> Us: What, do you have a crystal ball or something?
>
> Diana: No. I just know Ernie. He sees every reason in the world to lie, and no reason at all to tell the truth. Of course Ernie's going to lie! Don't you know Ernie? I'm not worried in the slightest.

The "restrictivists" amongst us might worry that, given the truth of this reply, Ernie's freedom not to lie is indeed called into question: his hardened character removes his freedom.[6] But I wish simply to sidestep this issue. For this is not the reply that we receive; we are not assuming that Ernie's lying is – in the perfectly ordinary way – predictable on the basis of his character. If it *isn't*, however, and Ernie is indeed torn, then whence comes Diana's utter certainty that he will lie? Again: she should be nervous. It is nerve-racking that she isn't.

A methodological interlude

[6] For a defense of restrictivism, see van Inwagen 1989. For criticism, see Fischer 1994: 47 - 62.

Such is, in brief, my proposed manipulation argument against classical compatibilism. But here it is worth pausing to make a point about the dialectical role of the "Diana scenario" in the relevant argument. To repeat, my contention is this: it is in part because Diana is not nervous that we have reason to suppose that Ernie is not free. Here it is *tempting* instead to write: it is because Diana is not nervous that Ernie is not free. Putting my point in this way may be tempting. However, this temptation, if any are subject to it, should be resisted. I am *not* claiming that it is because Diana is not nervous that Ernie *is not* free. Rather, and crucially, I am claiming that it is because Diana is not nervous that we can *tell* that Ernie is not free – or at any rate, Diana's not being nervous can *help us* in telling that Ernie is not free. Diana's not being nervous is good *epistemic reason* to suppose that Ernie is not free. Needless to say, Diana's not being nervous is not, on my proposal, part of what *brings it about* or *makes it the case* that Ernie is not free. It is, again, simply an *indicator* of Ernie's unfreedom.

We can look at the matter this way. Suppose I said (what I do not say, and do not think should be said) that it is because Diana is not nervous that Ernie is not free. Such a position can *seem* to indicate the following: if the scenario were otherwise perfectly similar concerning Ernie, but Diana *was* nervous, then Ernie *would* (or may) be free. But this is quite simply a mistake. Suppose that Diana proceeds as in the above scenario, thereby in fact determining that in 30 years Ernie shall lie. But suppose that, after a while, Diana forgets what she's done concerning Ernie – did she calculate correctly? Was the law she put in place indeed the deterministic one? She forgets – but now it is too late, and she must sit back and watch and see what Ernie does. Accordingly, she becomes nervous regarding whether he will in fact lie or tell the truth. Now, plainly, Diana's being nervous, on my view, could not possibly *bring it about* that Ernie is in fact free to do otherwise. If Ernie were not free to do otherwise when Diana remembers what she's done (correctly determine for Ernie to lie), then plainly he is *also* not free when Diana has set up precisely the same determining conditions, but has simply forgotten that she's done so. Her forgetting, in itself, plainly could make no difference regarding Ernie's freedom. However, if it were *because* she is not nervous that Ernie is not free, then her forgetting may do precisely that – namely, engender some nervousness, thereby preventing her lack of nervousness from causing Ernie not to be free. But this is plainly a mistake.

Now, does this realization – that Diana's not being nervous does not and could not *make* Ernie unfree – undermine the force of the argument? No. For Diana's not being nervous could play a role in helping us *see* that Ernie is not free, without playing a role in *making Ernie unfree.* That is, the argument trades on tapping into (or trying to tap into) something like a pre-theoretical judgment (or a connected web of judgments) we have about the freedom to do otherwise, viz., that freedom to do otherwise should imply Diana's being nervous. That, the proponent of the argument hopes you'll agree, is part of our pre-theoretical understanding of freedom to do otherwise. (If it isn't, then the argument is not a good argument, and that's that.) If Ernie is indeed free to do otherwise, then why isn't Diana nervous? Her non-nervousness, again, is a reliable indicator of Ernie's non-freedom. Insofar as we *do* judge that freedom to do to otherwise should imply Diana's being nervous, this strategy is perfectly sensible, and perfectly legitimate – even if we concede that her nervousness is, in the sense described above, *irrelevant* to Ernie's freedom. In a word, Diana's non-nervousness can play an *epistemic* role in indicating something about Ernie's freedom, without playing a *causal* or (perhaps) *metaphysical* role in Ernie's freedom. And this is all that matters for the cogency of the argument.[7]

The Consequence Argument and the Direct Argument

At this stage, I wish to situate the argument I have presented within the context of two traditional arguments in the debates about free will and moral responsibility (with which I assume broad familiarity): the Consequence Argument, and the Direct Argument.[8] By way of review: the Consequence Argument is an argument for the incompatibility of free will (understood as the ability to do otherwise) and determinism. The central claim

[7] Here I am anticipating responses to the manipulation argument along the lines of Fischer 2011 and King 2013. Their challenge can be summarized as follows:

> Either the manipulation does work in making the agent unfree, in which case the manipulation scenario is after all not equivalent to natural determination, or the manipulation does no work in making the agent unfree, in which case it is irrelevant to moral responsibility, and manipulation arguments get us nowhere.

As I am trying to indicate, the problem with this argument regards the second 'horn' of the alleged dilemma. As I have explained, the proponent of the manipulation argument can, and clearly should, grant that that the manipulation does not *make* the agent unfree, but simply *shows* or *indicates* something about the agent's freedom. For more on this point, see Todd 2013. For a defense of the manipulation argument's propriety (with which I basically agree), see Tognazzini 2014.

[8] See van Inwagen 1983: 55 – 105 (for the former) and 183 – 189 (for the latter).

undergirding the argument is as follows: If you are powerless to prevent one thing, and there is a necessary connection between that thing and another, then you are powerless to prevent the second thing as well. More particularly, powerlessness "transfers" across necessary connections: this is the transfer of powerlessness principle. Since, on determinism, there is a necessary connection between the past and the laws of nature with any fact at all, we therefore move from our powerlessness over the past and the laws to our powerlessness over *anything*.[9] Accordingly, determinism implies that no one has the power the prevent anything that he or she in fact does – which is other words for the claim that no one has the freedom to do otherwise.

Now, the Consequence Argument simply targets the freedom to do otherwise. In order to forge a connection to *moral responsibility*, we would have to suppose, in accordance with (some suitable version of) the principle of alternative possibilities, that such freedom is required for responsibility. Accordingly, the Consequence Argument targets responsibility only *indirectly* – *via* a conjunction with a principle like PAP. The Direct Argument, however, uses similar "ingredients" as the Consequence Argument, but it targets moral responsibility *directly*. The central claim undergirding the argument is as follows: If you are not morally responsible for one thing, and there is a necessary connection between that thing and another, then you are not morally responsible for the second thing either. More particularly, non-responsibility "transfers" across necessary connections: this is the transfer of non-responsibility principle. As before, given determinism, we thus transfer our non-responsibility for the past and the laws to our non-responsibility for *anything*. Accordingly, determinism implies that no one is morally responsible for anything that he or she does.

Now we may note the following. It is, it would appear, widely agreed that the Consequence Argument is *more plausible* than the Direct Argument. That is, it is widely agreed that the transfer of *powerlessness* principle is more evident than the parallel transfer of *non-responsibility* principle. Indeed, if I may sum up the current "state of the debate" concerning these arguments, we may note the following. There is, at present, disagreement concerning the validity of the transfer of non-responsibility principle; cases of "simultaneous overdetermination" would seem to provide counterexamples to the principle in question, counterexamples *both* sides of the debate would seem to

[9] It is controversial whether the "Consequence Argument" indeed does require the transfer of powerlessness principle; Fischer 1994: 62 – 66, for instance, contends that it doesn't. I hereby set this question aside.

recognize.[10] At any rate, such cases, many seem to think, call into question the dialectical usefulness of an appeal to the transfer of non-responsibility principle; in particular, if the principle is modified so as to escape such cases, then it becomes dangerously close simply to the assertion of incompatibilism.[11] However, there is *not* a similar state of affairs concerning the parallel transfer of *powerlessness* principle. (Here I am thinking of what has come to be called "Beta-box".[12]) There are *no* cases that are currently widely agreed (by partisans of both sides of the debate) to call *this* principle into question. And there have emerged no cases that have forced a modification of *this* principle, such that any modified principle appears to be lacking in dialectical force.[13]

Here, then, I wish to make several points. First: the above state of affairs accords with my own intuitive judgment: it is more plausible that Ernie lacks *freedom to do otherwise* than that he is simply *not responsible*. Now, I believe that it *is* plausible that Ernie is not responsible – and I believe that this is plausible, just on considering the facts of the story, and *not* only in conjunction with a principle like PAP. As I see it, the fact that Ernie is guaranteed to live out Diana's plan for his life, by itself, *indicates* something about his responsibility – namely, that he lacks it. That is, in just the same way that Diana's non-nervousness (to me) *indicates* his non-freedom, the fact that his life is entirely scripted in advance by someone else (to me) indicates his non-responsibility. It is, I think, simply part of our (or anyway my own) pre-theoretical "stock of judgments" about moral responsibility that a morally responsible agent's life cannot be entirely "scripted out" in the relevant way. This is, anyway, how I see it – and how many others see it.[14] Nevertheless, the point remains: *however* directly intuitive it is (on account of Ernie's living out Diana's plans) that Ernie is simply not responsible, it is even *more* directly intuitive (on account of Diana's non-nervousness) that he lacks freedom to do otherwise. In particular, I suspect that at least *some* who judge that Ernie is (or may be) *responsible* will nevertheless judge that Ernie is *not* free to do otherwise – and I suspect that *no one* who judges that Ernie is indeed free to do otherwise will nevertheless judge that Ernie is not

---

[10] For such counterexamples, see Fischer and Stump 2000 and Fischer 2004 (both reprinted in Fischer 2006).

[11] See Fischer 2004 in Fischer 2006: 168 – 173.

[12] Beta-box can be rendered as follows: Np, $\Box(p \rightarrow q) \vdash Nq$, where "N" is the "no one has a choice about" operator, and "$\Box$" is broadly-logical (or metaphysical) necessity. Beta-box is so-called because it replaces an "N" in van Inwagen's original "principle Beta" ($Np, N(p \rightarrow q) \vdash Nq$) with the "box" of metaphysical necessity.

[13] See, e.g. Widerker 1987, Finch and Warfield 1998, van Inwagen 2000, and Speak 2011: 120. As Campbell (2011: 49) writes, "There are no known counterexamples to principle (β') [Beta-box]."

[14] For a defense of this point of view, see Todd 2011, 2012, and 2013.

responsible.[15] The plausibility of Ernie's not being free to do otherwise highlights, I believe, the importance of the denial of PAP for compatibilist theories of moral responsibility.

But now we must return to the Consequence Argument and the Direct Argument. As I had just explained, it is widely agreed that the Consequence Argument is more plausible than the Direct Argument: whereas there remains controversy regarding the transfer of non-responsibility principle and its dialectical propriety, there is *not* a parallel controversy regarding the transfer of powerlessness principle. And this fact could seem to call into question the very purpose of this paper. After all, the primary point of this paper is to offer an argument against classical compatibilism. By the looks of it, however, it can seem that this project is unnecessary, for, one may think, the incompatibility of the ability to do otherwise and determinism has in effect already been established – or established to nearly everyone's satisfaction. As I noted, and as nearly everyone agrees, there are no good counterexamples to Beta-box. What more needs to be seen?

Such a line of thought, however, misses what is arguably the most attractive "reply" to the Consequence Argument. The most attractive reply to the argument does not attempt to construct a counterexample to "Beta-box" that will sway even the incompatibilist. (Nor does it maintain that we can prevent the past or the laws of nature.) Rather, the most attractive reply to the Consequence Argument seeks to show that the transfer principle is valid only for a sense of "power to do otherwise" that simply does not appear in ordinary life and in our ordinary thought and talk. When we *ordinarily* say that someone could have done otherwise, and when we *ordinarily* think of ourselves as free to do otherwise, then we have in mind a sense of "power" for which the transfer principle is simply invalid. All we mean is something like the following: *if* that person had wanted (or had tried, or…) to do otherwise, nothing would have prevented her from doing so. Accordingly, she *could have* done otherwise. However, once the "conditional" nature of what we ordinarily mean by "freedom to do otherwise" is understood, we'll see that the transfer principle is invalid as applied to what we ordinarily mean.[16]

---

[15] This point accords with Widerker's contention that the Direct Argument depends on the Consequence Argument. If you think that a determined agent may have freedom to do otherwise, you will *not* think that she must lack moral responsibility. For discussion, see Widerker 2002.

[16] This point has been made in various places. Kane, e.g., makes the point forcefully (and simply) in his 2005: 27 - 28.

The best way to deflect this compatibilist reply to the Consequence Argument is by *embedding* the incompatibilist's analysis of the "freedom to otherwise" within a wider set of judgments and attitudes we pre-theoretically have about such freedom. And this is precisely what the manipulation argument seeks to do: it seeks to *situate* our *ordinary* judgments of someone's being free to otherwise in a context in which it can then be seen that the ordinary assumption of such freedom implies the falsity of determinism. It is in this sense that the manipulation argument can *support* the Consequence Argument.[17] Recall the (modified) Diana scenario I presented above. We initially assume that Ernie is free not to lie, as Diana wants – free, precisely in the ordinary sense of "free not to lie". We assume it is up to Ernie whether to lie – up to, precisely in the ordinary sense of its being "up to someone" whether to lie. No special context has yet been created. And yet it seems clear that, if Ernie is free in this sense, then Diana should be nervous – absent some appeal to Diana's possessing a crystal ball. And once we understand the grounds for Diana's non-nervousness, precisely the ordinary sense of freedom we thought Ernie had is now called into question.[18]

The upshot is the following. The Consequence Argument is, by itself, a powerful argument. And perhaps that argument, by itself, is enough to overturn classical compatibilism and establish the incompatibility of freedom to do otherwise and determinism. Perhaps. Even if this were so, however, this would not imply that there could not be *additional* evidential considerations that support the denial of classical compatibilism. Indeed, if the Consequence Argument is sound, and sound *as applied to our ordinary conception* of freedom to do otherwise, then it would be frankly bizarre if its

---

[17] Essentially the same point may be made concerning manipulation arguments (that directly target moral responsibility) and the Direct Argument: the two arguments are mutually supporting and reinforcing. Indeed, this point has been made by Michael McKenna: "This kind of argument, a Manipulation Argument, could be used in the service of advancing the Direct Argument, and in particular, in the service of arguing for the controversial claims of nonresponsibility that I have focused upon in my rejection of the Direct Argument." (2008: 380)

[18] As an anonymous referee for this journal has noted, this does not imply that all senses of "having an ability" are incompatible with determinism. The incompatibilist should allow that the possession of *general abilities* is compatible with determinism; we may allow that Ernie possessed the *general ability* to tell the truth (he knows what telling the truth is, at least sometimes tells the truth, and so forth). As Randolph Clarke notes (following others), there are various things we might have in mind when saying that *S* "has an ability to (is able to, can) *A*", one of which is simply that "*S* has a general capacity to *A*." But we might also have in mind the more specific, "*S* has a general capacity to *A*, and *S* has a choice (at some specified time) about whether she (at some specified time) exercises that capacity." (337) As Clarke notes, "Something of the sort characterized by [this latter claim] is what many writing on free will have in mind." (339). It is this claim that I have targeted in this paper.

soundness did not "ramify out" into our wider thought and talk about freedom to do otherwise. That is, it would be bizarre if there were no *further* evidence (buried in our dispositions to ascribe or deny freedom) that freedom to do otherwise is incompatible with determinism. Indeed, we would expect us to have a stock of judgments (roughly speaking) that support the incompatibilist's analysis of freedom. And this is, indeed, precisely what I have tried to show that we do in fact find. Further, if we did *not* find the manipulation argument at all compelling, this would be excellent evidence that the Consequence Argument is leading us astray. The manipulation argument, however, *is* compelling – or so it seems to me. The two arguments are therefore mutually supporting and reinforcing.

Conclusion

In this paper, I have primarily advanced a manipulation argument for the incompatibility of freedom to do otherwise and determinism. As I see it, if Diana genuinely gave Ernie free will, then she should be nervous that he won't use it precisely as she would like. This thought, even if not absolutely irresistible, nevertheless seems deeply compelling. If so, then we have further, additional reason for the following conclusion: classical compatibilists should be nervous.[19]

References

Campbell, Joseph K. 2011. *Free Will.* Malden, MA: Polity.

Capes, Justin. 2014. "The Flicker of Freedom: A Reply to Stump," *Journal of Ethics* 18: 427 – 435.

Cohen, Yishai. 2015. "The Manipulation Argument, at the Very Least, Undermines Classical Compatibilism," *Philosophia* 43: 291 – 307.

Clarke, Randolph. 2009. "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism," *Mind* 118: 323-351.

Finch, Alicia and Ted Warfield. 1998. "The Mind Argument and Libertarianism," *Mind* 107: 515 – 528.

Fischer, John Martin. 2004. "The Transfer of Non-Responsibility," in J. Campbell, M. O'Rourke, D. Shier, eds., *Freedom and Determinism: Topics in Contemporary Philosophy*

[19] For comments on previous drafts of this paper, I wish to thank John Martin Fischer, Neal Tognazzini, Al Mele, Andrew Bailey, and Derk Pereboom. I presented versions of this paper in October 2015 at Fordham University and Rutgers University; thanks to audiences at both events, especially Amy Seymour and Joe Vukov at Fordham, and Philip Swenson, Andrew Moon, Dean Zimmerman, Holly Smith, David Black, Brian Cutter, Pamela Robinson, and Eli Shupe at Rutgers.

Series 2 (MIT Press): 189 – 209.

Fischer, John Martin. 2006. *My Way: Essays on Moral Responsibility*. Oxford: Oxford University Press.

Fischer, John Martin. 2011. "The Zygote Argument Remixed," *Analysis* 71: 267 – 272.

Fischer, John Martin and Eleonore Stump. 2000. "Transfer Principles and Moral Responsibility," *Philosophical Perspectives* 14: 47 – 56.

Franklin, Christopher Evan. 2011. "Neo-Frankfurtians and Buffer Cases: The New Challenge to the Principle of Alternative Possibilities," *Philosophical Studies* 152: 189–207.

Kane, Robert. 2005. *A Contemporary Introduction to Free Will*. Oxford: Oxford University Press.

King, Matt. 2013. "The Problem with Manipulation," *Ethics* 124: 65 – 83.

McKenna, Michael. 2008. "Saying Goodbye to the Direct Argument the Right Way." *Philosophical Review* 117.3: 349-383.

Mele, Alfred. 2006. *Free Will and Luck*. Oxford: Oxford University Press.

Pereboom, Derk. 2014. *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.

Speak, Daniel. 2011. "The Consequence Argument Revisited," in *The Oxford Handbook of Free Will*, ed. Robert Kane, pp 115 – 130.

Swenson, Philip. 2015. "A Challenge for Frankfurt-Style Compatibilists," *Philosophical Studies* 172: 1279 – 1285.

Todd, Patrick. 2011. "A New Approach to Manipulation Arguments," *Philosophical Studies* 152: 127 – 133.

Todd, Patrick. 2012. "Manipulation and Moral Standing: An Argument for Incompatibilism," *Philosophers' Imprint* 12: 1 – 18.

Todd, Patrick. 2013. "Defending (a modified version of) the Zygote Argument," *Philosophical Studies* 164: 189 – 203.

Tognazzini, Neal. 2014. "The Structure of a Manipulation Argument," *Ethics* 124: 358 – 369.

Van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Oxford University Press.

Van Inwagen, Peter. 1989. "When is the will free?" *Philosophical Perspectives* 3: 76 – 94.

Van Inwagen, Peter. 2000. "Free Will Remains a Mystery," *Philosophical Perspectives* 14: 1 – 19.

Widerker, D. 2002. "Farewell to the Direct Argument," *Journal of Philosophy* 99:316 - 324.